Published in final edited form as:

J Biotechnol. 2011 December 10; 156(3): 188–196. doi:10.1016/j.jbiotec.2011.08.013.

Response surface optimization of the heparosan *N*-deacetylation in producing bioengineered heparin

Zhenyu Wang^{a,e}, Jennifer Li^a, Samantha Cheong^a, Ujjwal Bhaskar^{c,e}, Onishi Akihiro^e, Fuming Zhang^{b,e}, Jonathan S. Dordick^{a,c,d,e}, and Robert J. Linhardt^{a,b,c,d,e}

^aDepartment of Biology, Rensselaer Polytechnic Institute; 110 8th Street, Troy, NY 12180, USA

^bDepartment of Chemistry and Chemical Biology, Rensselaer Polytechnic Institute; 110 8th Street, Troy, NY 12180, USA

^cDepartment of Chemical and Biological Engineering, Rensselaer Polytechnic Institute; 110 8th Street, Troy, NY 12180, USA

^dDepartment of Biomedical Engineering, Rensselaer Polytechnic Institute; 110 8th Street, Troy, NY 12180, USA

^eCenter for Biotechnology and Interdisciplinary Studies, Rensselaer Polytechnic Institute; 110 8th Street, Troy, NY 12180, USA

Abstract

The chemical step in the chemoenzymatic synthesis of bioengineered heparin has been examined and optimized statistically using a response surface methodology. A four factor, two level full factorial design experiment and a three factor Box-Behnken design were carried out. The goal was to establish a method to prepare *N*-sulfo, *N*-acetyl heparosan of the desired *N*-acetyl content, number average molecular weight, and in maximum yield by controlling the reactant concentrations, reaction time and reaction temperature. The response surface models obtained were used to predict the reaction conditions required to optimally prepare *N*-sulfo, *N*-acetyl heparosan from *E. coli* generated heparosan starting material of different molecular weights.

Introduction

Heparin is one of the most widely used anticoagulant drugs (Dahlback, 2000; Linhardt, 1991). Its anticoagulant activity is exploited in circumstances where the normal propensity for blood to clot must be overcome. Surgical procedures often require heparin, as do extra-corporeal therapies, such as heart-lung oxygenation and kidney dialysis. Heparin is also coated on the surface of experimental or medical devices such as test tubes and renal dialysis machines to form an anticoagulant surface. Heparin also controls the clotting of blood in diseased vessels, and is therefore used to treat deep vein thrombosis and acute coronary syndrome (Agnelli et al., 1998; Linhardt, 1991).

Currently, United States Pharmacopeial (USP) heparin is primarily extracted from porcine intestine. The heparin preparation process consists of several steps, not following current good manufacturing practice (cGMP), to extract raw heparin from porcine intestine that are

© 2011 Elsevier B.V. All rights reserved.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

usually performed at or near the slaughterhouse (Okuyama et al., 1975). The resulting raw heparin then goes through final purification steps in a pharmaceutical company under cGMP to afford USP heparin (Gerard and Pierre, 1961; Vidic, 1981). The current production process for pharmaceutical heparin for qualification as USP heparin has several drawbacks: (1) the animal origin of heparin results in the risk of animal virus and prion impurities being present; (2) non-cGMP process steps increases the risk of product contamination or adulteration; and (3) the amount of heparin available is limited by the amount of porcine intestine available, the major reason that the U.S. relies on foreign countries to supply heparin raw material. A heparin contamination crisis occurred in early 2008, which was marked by the increase in serious adverse events associated with heparin therapy and affecting thousands of patients (Kishimoto et al., 2008; Liu et al., 2009). The contaminant was later identified as oversulfated chondroitin sulfate (OSCS) (Guerrini et al., 2008; Liu et al., 2009). While it remains unclear exactly how OSCS was introduced into the heparin products, it appears that an adulteration took place during the preparation or consolidation of raw heparin prior to the cGMP process.

Efforts have been made to eliminate the problems associated with animal-sourced heparin production. A number of publications have proposed using a combination of fermentation, chemical, and/or enzymatic reactions to make heparin analogs (Casu et al., 1994; Chen et al., 2005; Chen et al., 2007; Kuberan et al., 2003a; Kuberan et al., 2003b; Lindahl et al., 2005; Linhardt Robert J., 1997; Liu et al., 2009; Zhang et al., 2008). In 2009, our laboratory initiated a research consortium to develop a commercially viable process to produce kilogram quantities of bioengineered (non-animal sourced) heparin within a period of five years. This bioengineered heparin would be chemically and biologically equivalent to pharmaceutical porcine heparin and utilize an industrially scalable process to produce a generic version of heparin. Our previous studies have indicated the bioengineered heparin we prepared have comparable chemical structure, antithrombin binding activity and *in vitro* anticoagulant activity with USP porcine heparin (Wang et al., 2011b; Zhang et al., 2008).

Bioengineered heparin preparation starts with the fermentation of *E. coli* K5 to afford heparosan polysaccharide for subsequent chemoenzymatic modification (Zhang et al., 2008). The fermentation process has been extensively studied and controlled to yield ideal heparosan material as the precursor for preparing bioengineered heparin, and the fermentation process has been optimized to increase the yield and lower the cost (Wang et al., 2010). The next step in the bioengineered heparin process is the chemical *N*-deacetylation and *N*-sulfonation step (Figure 1). To enable the final bioengineered heparin resemble the structure of the USP porcine heparin, the ratio of *N*-acetyl glucosamine to *N*-sulfo glucosamine must match those in porcine heparin. This requires fine control of the chemical *N*-deacetylation reaction to preserve the appropriate proportion of *N*-acetyl groups. This is a critical step for several reasons: (1) the *N*-deacetylation reaction directly affects the *N*-acetyl content of the bioengineered heparin produced; (2) chemical *N*-deacetylation uses aqueous sodium hydroxide that reduces the molecular weight of heparosan through a limited depolymerization reaction, affording a smaller polysaccharide chain comparable in size to heparin; and (3) the *N*-sulfo/*N*-acetyl heparosan obtained following *N*-sulfonation serves as the new backbone for subsequent enzymatic modifications, and the positioning of the *N*-sulfo and *N*-acetyl groups directly impacts the activity of the enzymes in affording the desired final structure of heparin (Chen et al., 2005; Kuberan et al., 2003a; Kuberan et al., 2003b).

Commercial USP heparin samples were reported to have an average *N*-acetyl content of 14.8%, with different USP heparin samples ranging from 11.9% to 17.6% (Guerrini et al., 2001). Heparin is polydisperse, consisting of polysaccharide chains of range of different sizes. We examined the average molecular weight of commercial heparins by Size Exclusion

Chromatography (SEC), which had an average number average molecular weight (M_N) of 15.1 KDa and a M_N range of from 14.3 KDa to 15.9 KDa. For the bioengineered heparin to be chemically equivalent to current USP heparins, the *N*-acetyl, *N*-sulfo heparosan intermediate obtained by chemical *N*-deacetylation and *N*-sulfonation (Figure 1) must match USP heparins with *N*-acetyl content and M_N . The resultant *N*-acetyl, *N*-sulfo heparosan intermediate then would be processed through enzymatic modification steps into a bioengineered heparin closely resembling USP heparin.

The chemical *N*-deacetylation of heparosan with aqueous sodium hydroxide has two major effects on the heparosan chain precursor. First, the *N*-acetyl groups of the *N*-acetylglucosamine residue are partially (or completely) removed to afford unsubstituted amino groups, and second, the heparosan polysaccharide chain is partially depolymerized through a β -elimination mechanism (Jandik et al., 1996), reducing its molecular weight. There are four major factors in the *N*-deacetylation step that might impact the product's *N*-acetyl content, average molecular weight and recovery yield. These are the initial concentration of heparosan, the NaOH concentration, the reaction time, and the reaction temperature. By controlling these four factors, one may be able to obtain a product with desired *N*-acetyl content, average molecular weight and a high recovery yield.

The heparosan reactant, produced from *E. coli* K5 fermentation, varies in M_N due to variation in the medium composition and culture conditions (Manzoni et al., 2000; Wang et al., 2010). Even under a well-controlled fermentation operation, the heparosan M_N may vary from batch to batch due to the complexity of bioprocesses, and the expression of a K5 lyase, which depolymerizes the heparosan chain through a β -elimination mechanism. In batch culture with a culture time around 14 h, the M_N of the recovered heparosan is usually around 50 KDa (Wang et al., 2010). When the heparosan production is carried out with a fed-batch mode in a fermentor, the prolonged fermentation time increases the action of the K5 lyase, thus, decreasing the M_N of heparosan produced (Wang et al., 2011a; Wang et al., 2010). A different degree of depolymerization is required, based on the different starting heparosan M_N , to produce bioengineered heparin of the M_N resembling USP heparin. This may be achieved by varying the *N*-deacetylation reaction conditions. However, the reaction product must also meet another criterion, the appropriate *N*-acetyl content, and a third objective, a maximized product yield. Thus, the goal of the heparosan *N*-deacetylation becomes a multi-objective optimization problem, in which three objectives are targeted: *N*-acetyl content, M_N , and product yield. A model is necessary in solving this complex problem, and will be especially useful in predicting the reaction conditions for producing the desired bioengineered heparin product from heparosan produced in different *E. coli* K5 fermentation batches in the expected industrial production of bioengineered heparin.

Response surface methodology (RSM) is a collection of statistical and mathematical techniques useful for developing, improving, and optimizing processes (Myers and Montgomery, 2002). The most extensive applications of RSM are in situations where several input variables potentially influence some performance measure or quality characteristic of the product or process (Myers and Montgomery, 2002). This performance measure or quality characteristic is called the response. For example, RSM has been used to optimize fermentation medium (Barrington and Kim, 2008; Ma et al., 2009), improve analytical techniques (Ferreira et al., 2007), and modeling coal grinding (Aslan and Cebeci, 2006). Box-Behnken designs (BBD) are one class of the experimental designs for response surface methodology. They are rotatable or nearly rotatable second-order designs based on three-level incomplete factorial designs (Box and Behnken, 1960). For three factor BBD, its graphical representation can be seen as a cube that consists of the central point and the middle points of the edges (Ferreira et al., 2007; Myers and Montgomery, 2002). The number of experiments (N) required for the development of BBD is defined as $N=2k(k-1)$

+ C_0 , (where k is number of factors and C_0 is the number of central points). A comparison between the BBD and other response surface designs (central composite, Doehlert matrix and three-level full factorial design) has demonstrated that the BBD and Doehlert matrix are slightly more efficient than the central composite design but much more efficient than the three-level full factorial designs where the efficiency of one experimental design is defined as the number of coefficients in the estimated model divided by the number of experiments (Ferreira et al., 2007). Another advantage of the BBD is that it does not contain combinations for which all factors are simultaneously at their highest or lowest levels. So these designs are useful in avoiding experiments performed under extreme conditions, for which unsatisfactory results might occur. Conversely, they are not indicated for situations in which we would like to know the responses at the extremes, that is, at the vertices of the cube (Ferreira et al., 2007; Myers and Montgomery, 2002).

In this study, we utilize a response surface method to solve a multi-objective optimization problem associated with the *N*-deacetylation of heparosan. Reaction conditions for different number average molecular weight heparosan are predicted with the model developed. The products of the *N*-deacetylation obtained with the predicted reaction conditions met our expectation targets.

Materials and Methods

Preparation of *Escherichia coli* K5 heparosan

E. coli K5 heparosan was produced by *E. coli* K5 strain (ATCC #23506) fermentation and purified from the culture supernatant as described previously (Wang et al., 2011a; Wang et al., 2010). More specifically, the K5 heparosan with M_N of 14.9 KDa and 14.0 KDa was produced by two different batches of exponential feeding fed-batch culture (Wang et al., 2010), and the K5 heparosan with M_N of 18.4 KDa was produced from a pH-stat fed-batch culture (Wang et al., 2011a). All the heparosan batches are purified with DEAE Sepharose Fast Flow anion exchange resin from GE Healthcare (Piscataway, NJ) (Wang et al., 2010).

N-Deacetylation/*N*-Sulfonation of K5 heparosan

K5 polysaccharide was dissolved in varying concentration of 1 ml NaOH solution, incubated at different temperatures for different lengths of time, the reaction mixture was then diluted to 5 ml, cooled on ice, and adjusted to pH 7 with HCl. Sodium carbonate (60 mg) and trimethylamine-sulfur trioxide complex (60 mg) were added in a single step, and the mixture was incubated for 12 h. An equal portion of sodium carbonate and trimethylamine-sulfur trioxide was again added after 12 h and the chemoselective *N*-sulfonation was continued for an additional 12 h at 50 °C. The solutions were then brought to room temperature, dialyzed overnight against distilled water using a 3500 Da molecular weight cut-off (MWCO) cellulose membrane. Commercial heparin typically has a M_N around 15 KDa, with very few chains below 3500 Da, thus, the use of a 3500 Da MWCO cellulose membrane does not affect the product chains that are within our target range, while can effectively remove salt from the sample. The dialyzate was lyophilized to obtain salt-free, *N*-sulfo, *N*-acetyl heparosan polysaccharide (Kuberan et al., 2003a). The starting heparosan material used for the factorial design and Box-Behnken design had a M_N of 14.9 KDa.

NMR analysis of *N*-sulfo, *N*-acetyl heparosan

^1H -NMR was conducted on a Bruker 600 MHz NMR spectrometer. The samples were prepared in 5 mm standard NMR tubes after lyophilization in D_2O . Acquisition of the spectra was carried out using TOPSPIN 2.0 software. All the spectra were acquired at the temperature of 298 K. A recycle delay time of 10 s was used. The acquired ^1H -NMR spectra were processed with Mnova NMR software for phase and baseline correction. The H1

proton of the glucosamine was used for quantifying the *N*-acetyl content, as it shows a peak at around 5.31 ppm when *N*-acetyl glucosamine and shifts to around 5.55 ppm when *N*-sulfo glucosamine. A typical ^1H -NMR spectrum of *N*-sulfo, *N*-acetyl heparosan produced is shown in Figure 2 with the peaks assigned. *N*-acetyl content was calculated as the peak area of *N*-acetyl glucosamine H1 proton divided by the sum of peak areas of *N*-acetyl glucosamine H1 proton and *N*-sulfo glucosamine H1 proton. The peak areas were calculated using the “manual integration” or “line fitting” function of Mnova NMR software.

Size exclusion chromatography of K5 heparosan and *N*-sulfo, *N*-acetyl heparosan for molecular weight determination

SEC was performed using TSK-GEL G3000PWxl or G4000PWxl size exclusion column with a sample injection volume of 20 μL and a flow rate of 0.6 ml/min on an apparatus composed of a Shimadzu LC-10Ai pump, a Shimadzu CBM-20A controller and a Shimadzu RID-10A refractive index detector. The mobile phase consisted of 0.1 M NaNO_3 . The column was maintained at 40°C with an Eppendorf column heater during the chromatography. The SEC chromatograms were recorded with the Lcsolution version 1.25 software and analyzed with its “GPC Postrun” function. For molecular weight determination of K5 heparosan, TSK-GEL G4000PWxl size exclusion column was used, and hyaluronan standards of different molecular weights (30.6 kDa, 54 kDa, 125 kDa and 250 kDa), purchased from Hyalose L.L.C. (Oklahoma City, Oklahoma), were used as calibrants for the standard curve. For molecular weight determination of *N*-sulfo, *N*-acetyl heparosan, TSK-GEL G3000PWxl size exclusion column was used, and heparin oligosaccharides of different molecular weights (dp 6, dp 10, dp 16 and dp 20) purchased from Iduron (Manchester, UK) were used as calibrants for the standard curve.

Experimental design

Full factorial design

A four factor, two-level full factorial design was generated with the Minitab software from the tab “Stat -> DOE -> Factorial -> Create Factorial Design” and performed with the four factors being heparosan concentration, NaOH concentration, reaction time and reaction temperature. Each of the factors has two coded levels, and the corresponding uncoded values are illustrated in Table 1a. The full factorial design was based on the first-order model:

$$Y = \beta_0 + \sum \beta_i X_i \quad (1)$$

where, Y is the response, β_0 is the model intercept and β_i is the linear coefficient, and X_i is the level of the independent variable. The purpose of the full factorial design was to identify the factors that significantly affect the product properties in terms of *N*-acetyl content, M_N and yield, and further investigate their effects quantitatively for predicting the reaction conditions in the subsequent response surface design. Another value of the full factorial design is to identify the insignificant factors and exclude them from the next stage of experimental design, thus reducing the number of runs needed for the next response surface design. These insignificant factors can be set to a level that favors the process economics to benefit the real production process.

The full factorial design was carried out in duplicate to be statistically reliable. The experimental data were analyzed with the software Minitab 15 following the tab “Stat -> DOE -> Factorial -> Analyze Factorial Design”.

Box-Behnken design

A Box-Behnken design was generated with Minitab software from the tab “Stat -> DOE -> Response Surface -> Create Response Surface Design” and carried out to quantitatively model the effect of the variables on the *N*-acetyl, *N*-sulfo heparosan product's *N*-acetyl content, M_N and yield. The levels of the three significant factors were redefined (Table 2a) to be in the vicinity of the center point and the practically operational region. The Box-Behnken design has 15 experimental runs with three runs at the center point (Table 2b) to develop response surface models that describe the relationship of reaction variables with the product's *N*-acetyl content, M_N and yield, respectively.

The 15 *N*-deacetylation experimental runs were conducted in 1 ml centrifuge tubes, and the reaction temperature was controlled with temperature-adjustable water bath. The experimental data was analyzed with MiniTab 15 following “Stat -> DOE -> Response Surface -> Analyze Response Surface Design” and fitted into a second-order equation. The quadratic equation model is as the following:

$$Y = \beta_0 + \sum \beta_i X_i + \sum \beta_{ii} X_i^2 + \sum \beta_{ij} X_i X_j \quad (2)$$

where Y is the predicted response; β_0 is the offset term; β_i is the linear effect; β_{ii} is the squared effect; β_{ij} is the interaction effect, and X_i and X_j are the dimensionless coded value of the variable x_i and x_j .

For statistical calculations, the relationships between the coded values and actual values are described by the following equation:

$$X_i = \frac{x_i - x_0}{\Delta x_i} \quad (3)$$

where X_i is the dimensionless coded value of the independent variable x_i ; x_i is the actual value of that independent variable; x_0 is the real value of the independent variable x_i at the center point and Δx_i is the step change.

Results

Full factorial design

Table 1b represents the conditions for running the four-factor, two level full factorial design and the responses. The high and low levels of the factors were chosen according to previous reported conditions for polysaccharide *N*-deacetylation (Erbing et al., 1976; Kuberan et al., 2003a). The design responses were analyzed with the Minitab software. Table 1c shows the first-order regression results and the significance levels of each factor. The P value level of 0.05 was used to determine whether a factor has significant effect on the responses. For *N*-acetyl content, factors X_2 (NaOH concentration), X_3 (reaction time) and X_4 (reaction temperature) all have a P value of less than 0.05, indicating they are significant in influencing the product *N*-acetyl content at the tested levels. X_1 (heparosan concentration) gives a P value of 0.946, greater than 0.05, indicating it does not have significant effect on product *N*-acetyl content. Similarly, X_2 (NaOH concentration), X_3 (reaction time) and X_4 (reaction temperature) all have significant effect on the depolymerization factor and product recovery, while the P values of X_1 (heparosan concentration) indicate heparosan concentration is not a significant factor in influencing the product depolymerization and recovery.

Because heparosan concentration was not significant in affecting product properties, it is set to its high level (10 g/L) to increase the process throughput capacity, so that more material could be processed in a single batch reaction for prospective commercial production. The regression analysis with Minitab fits the full factorial design data into the following first-order equations:

$$N\text{-acetyl content } Y_1 (\%) = 3.392 - 0.039 X_1 - 1.336 X_2 - 3.392 X_3 - 1.322 X_4 \quad (4)$$

$$\text{Depolymerization factor } Y_2 = 0.23 + 0.0137 X_1 - 0.0919 X_2 - 0.0484 X_3 - 0.1415 X_4 \quad (5)$$

$$\text{Product recovery } Y_3 (\%) = 39.75 + 0.47 X_1 - 8.5 X_2 - 4.01 X_3 - 31.54 X_4 \quad (6)$$

The negative coefficients of X_2 , X_3 and X_4 for all the three responses indicate they have negative effects on product *N*-acetyl content, depolymerization factor and product recovery.

Box-Behnken design experiment and response surface analysis

The effects of the three significant variables NaOH concentration, reaction time and reaction temperature were further analyzed with a Box-Behnken design, aiming to build a quadratic response surface model that describes the effects of these three variables on the product properties and can be used for predicting the reaction condition for heparosan with a different M_N from the heparosan used for the factorial design and Box-Behnken design.

Commercial heparins typically have a *N*-acetyl content of 11.9% – 17.6% (Guerrini et al., 2001), with an average *N*-acetyl content of 14.8% (Guerrini et al., 2001); and the commercial heparins have a M_N ranging from 14.3 KDa to 15.9 KDa, with an average M_N of 15.1 KDa. The *N*-sulfo, *N*-acetyl heparosan produced by heparosan *N*-deacetylation and *N*-sulfonation need to go through enzymatic C5 epimerization and *O*-sulfonations to become bioengineered heparin (Figure 1c). According to the C5 epimerization and *O*-sulfonation patterns, the targeted *N*-sulfo, *N*-acetyl heparosan should have a *N*-acetyl content of ideally 14.8%, or within the range of 11.9% – 17.6%; and the targeted M_N of the *N*-sulfo, *N*-acetyl heparosan should be ideally 11.7 KDa, or within the range of 11.0 KDa – 12.3 KDa. From the result of the full factorial design, the reaction condition of 2 M NaOH, reaction time 3 h, and reaction temperature 60 °C gives a product with properties that are very close to the targets. Thus, the reaction condition of 2 M NaOH, reaction time 3 h, and reaction temperature 60 °C was chosen as the center point for the Box-Behnken design. The high and low levels of each factor was chosen empirically (Table 2a) and all reactions were conducted at the 10 mg/ml heparosan concentration level, as heparosan concentration is not significant and a high level can increase process throughput. The Box-Behnken design and the responses are illustrated in Table 2b.

The Box-Behnken responses were analyzed with the Minitab software, and the regression result is illustrated in Table 2c. For the response product recovery? the P value for most of the variables and their quadratic terms are greater than 0.05, indicating they are not significant in affecting the product recovery within the experimental design range; only the term X_1^2 and intercept have P values of less than 0.05. However, the R^2 value is relatively low (0.823), indicating only 82.3% of the variability in the response can be explained by the model; moreover, the Analysis of Variance for the response “product recovery” with F test gives an F value of 2.59 and a P value of 0.154, indicating the regression model was not significant at the 95% significance level. Thus, the variability in the response “product

recovery” was considered to be caused by random error at the experimental design range and was not chosen as a subject for optimization.

The R^2 of the regression model for *N*-acetyl content is 0.995, and the *F* and *P* values were 119.87 and 0, respectively, indicating good fitting and significance of the regression model. Similarly, the regression model for depolymerization factor gave a R^2 of 0.960, *F* value of 13.47 and *P* value of 0.005, validating the fitting and significance of the regression model. The effects of the NaOH concentration, reaction time and reaction temperature on the product *N*-acetyl content and depolymerization factor can be explained with the following second-order polynomial equations:

$$\begin{aligned} N\text{-acetyl content } Y_1 (\%) = & 15.02 - 23.793X_1 \\ & - 8.083X_2 \\ & - 17.575X_3 \\ & + 10.527X_1^2 \\ & + 4.579X_2^2 \\ & + 5.973X_3^2 \\ & + 3.472X_1X_2 \\ & + 6.941X_1X_3 + 3.433X_2X_3 \end{aligned} \quad (7)$$

$$\begin{aligned} \text{Depolymerization factor } Y_2 = & 0.72963 - 0.03175X_1 \\ & - 0.09502X_2 \\ & - 0.2227X_3 \\ & + 0.01088X_1^2 \\ & + 0.02028X_2^2 \\ & + 0.01632X_3^2 \\ & - 0.0257X_1X_2 \\ & - 0.02684X_1X_3 - 0.04509X_2X_3 \end{aligned} \quad (8)$$

The effects of NaOH concentration, reaction time and reaction temperature on product *N*-acetyl content are illustrated with the three-dimensional response surface and contour plot in Figure 3a, b and c. The effects of NaOH concentration, reaction time and reaction temperature on product depolymerization factor are illustrated with the three-dimensional response surface and contour plot in Figure 4a, b and c. With the response surface model described with equation 7 and 8, the reaction condition that yield a product with both the right *N*-acetyl content and a specific depolymerization factor according to the starting heparosan material M_N can be solved.

The model was first tested on a K5 heparosan batch with a M_N of 14.0 KDa, which is smaller than the K5 material used to build the response surface model and would require a depolymerization factor of 0.83 to produce the *N*-sulfo, *N*-acetyl heparosan with the desired molecular weight of around 11.7 KDa. With the response surface model described by equation 7 and 8, a reaction condition in coded level of NaOH concentration at 0.535, reaction time at -1 and reaction temperature at 0.0777 should afford a *N*-sulfo, *N*-acetyl heparosan with *N*-acetyl content around 14.8% and M_N around 11.7 KDa. The reaction conditions correspond to 2.54 M NaOH concentration, 2 h reaction time and reaction temperature at 60.8 °C. The reaction was carried out at these conditions and the *N*-sulfo, *N*-acetyl heparosan product was analyzed to have a *N*-acetyl content of 15.6% and M_N of 11.3 KDa, close to the targets and within the commercial heparins' variability range. The model

was then used to predict the reaction conditions for another K5 heparosan batch, which was larger than the heparosan used to build the model having a M_N of 18.4 KDa. The response surface model predicts a coded reaction level of NaOH concentration at -0.238 , reaction time at 1, and reaction temperature at 0.143 to produce the *N*-sulfo, *N*-acetyl product with the desired *N*-acetyl content and M_N . The reaction conditions correspond to 1.76 M NaOH concentration, 4 h reaction time, and reaction temperature at 61.4 °C. The reaction was carried out with the above conditions and a *N*-sulfo, *N*-acetyl heparosan product with 13.5% *N*-acetyl content and 11.5 KDa M_N was obtained. The *N*-acetyl content and M_N are close to the target and within the commercial heparins' variability range.

Discussion

A response surface model was established by a full factorial design and a Box-Behnken design experiments. This model provides guidance in choosing the reaction conditions to obtain *N*-sulfo, *N*-acetyl heparosan with desired properties. Furthermore, the model is the basis for solving the multi-objective optimization problem in making the ideal *N*-sulfo, *N*-acetyl heparosan for bioengineered heparin production. A single reaction condition that meets both the criteria of the right *N*-acetyl content and M_N of the product can be obtained from the model equations.

The model equation 7 was developed to control the properties of the *N*-acetyl content of the *N*-sulfo, *N*-acetyl heparosan product. The starting heparosan material is 100% *N*-acetylated, the *N*-deacetylation step converts the *N*-acetyl group to *N*-amino group, and the *N*-sulfonation step completely sulfonates the *N*-amino group to become *N*-sulfo group. By varying reaction conditions described with equation 7, different remaining *N*-acetyl contents can be obtained in the *N*-deacetylation and *N*-sulfonation reactions. To prepare *N*-sulfo, *N*-acetyl heparosan to generate bioengineered heparin, 14.8% *N*-acetyl content is targeted, which is the reported average *N*-acetyl content of commercial heparins.

The model equation 8 was developed to control the product M_N . The term “depolymerization factor” was used to describe the extent of depolymerization occurred during the reaction. The depolymerization during the heparosan *N*-deacetylation was assumed to be random and not dependent on the starting material chain length. The extent of the depolymerization was assumed to be reaction condition dependent and, thus, can be varied by changing the reaction conditions to obtain different depolymerization extent for starting heparosan material with different M_N . These assumptions are theoretically reasonable and validated experimentally with our model. Thus, the model could be used to control the product M_N by varying the depolymerization factor.

Product recovery was originally included in the study with the full factorial design, where it is found to be significantly affected by NaOH concentration, reaction time and reaction temperature. However, when the variation range of the reaction condition factors was decreased in the Box-Behnken design, its variation was found to be likely random error and not significantly affected by the reaction condition factors. This observation can be explained with the dialysis step in the recovery process. The product after the reactions was dialyzed in a 3.5 KDa MWCO membrane. In the factorial design, some harsh reaction conditions (high NaOH concentration, high reaction temperature and long reaction time) caused extensive depolymerization of the polysaccharide and resulted in chain sizes corresponding to $M_N < 3.5$ KDa. These small chains could perfuse through the membrane and, thus, were not recovered. In the Box-Behnken design, the reaction conditions are comparatively mild, although different degrees of depolymerization occurred, very little polysaccharide was depolymerized to $M_N < 3.5$ KDa and, thus, essentially all the product could be recovered after dialysis.

The model converts the optimization of the heparosan *N*-deacetylation condition to the mathematical problem of solving the two equations 7 and 8, which can be easily achieved with Minitab and Matlab software. This was achieved with the Box-Behnken design, which allows an efficient estimation of the first-order and second-order coefficients with a relatively small number of experimental runs. Three factor Box-Behnken design has its best prediction ability when the predicted point falls within $\sqrt{2}$ radius from the center point (0,0,0) (Myers and Montgomery, 2002). In the case of this study, this range covers the heparosan starting material M_n with a lower limit of 12.6 KDa and upper limit of 24.6 KDa. The prediction variance may be large if the heparosan starting material has an M_n outside of this range.

Research Highlights

- Optimized chemical step in the chemoenzymatic synthesis of bioengineered heparin
- Optimized using a response surface methodology.
- A 4-factor, 2-level full factorial design experiment was used.
- A 3-factor Box-Behnken design was carried out.
- Established method to prepare *N*-sulfo, *N*-acetyl heparosan.

Acknowledgments

The authors are grateful for funding from the NIH in the form of grant # HL096972 and the support of the Bioengineered Heparin Consortium for supporting this research.

References

- Agnelli G, Piovella F, Buoncristiani P, Severi P, Pini M, D'Angelo A, Beltrametti C, Damiani M, Andrioli GC, Pugliese R, Iorio A, Brambilla G. Enoxaparin plus compression stockings compared with compression stockings alone in the prevention of venous thromboembolism after elective neurosurgery. *N Engl J Med*. 1998; 339:80–85. [PubMed: 9654538]
- Aslan N, Cebeci Y. Application of Box–Behnken design and response surface methodology for modeling of some Turkish coals. *Fuel*. 2006; 86:90–97.
- Barrington S, Kim JW. Response surface optimization of medium components for citric acid production by *Aspergillus niger* NRRL 567 grown in peat moss. *Bioresour Technol*. 2008; 99:368–377. [PubMed: 17267213]
- Box GE, Behnken DW. Some new three level designs for the study of quantitative variables. *Technometrics*. 1960; 2:455–475.
- Casu B, Grazioli G, Razi N, Guerrini M, Naggi A, Torri G, Oreste P, Tursi F, Zoppetti G, Lindahl U. Heparin-like compounds prepared by chemical modification of capsular polysaccharide from *E. coli* K5. *Carbohydr Res*. 1994; 263:271–284. [PubMed: 7805054]
- Chen J, Avci FY, Munoz EM, McDowell LM, Chen M, Pedersen LC, Zhang L, Linhardt RJ, Liu J. Enzymatic redesigning of biologically active heparan sulfate. *J Biol Chem*. 2005; 280:42817–42825. [PubMed: 16260789]
- Chen J, Jones CL, Liu J. Using an enzymatic combinatorial approach to identify anticoagulant heparan sulfate structures. *Chem Biol*. 2007; 14:986–993. [PubMed: 17884631]
- Dahlback B. Blood coagulation. *Lancet*. 2000; 355:1627–1632. [PubMed: 10821379]
- Erbing C, Granath K, Kenne L, Lindberg B. A new method for the *N*-deacetylation of carbohydrates. *Carbohydrate Research*. 1976; 47:C5–C7.

- Ferreira SL, Bruns RE, Ferreira HS, Matos GD, David JM, Brandao GC, da Silva EG, Portugal LA, dos Reis PS, Souza AS, dos Santos WN. Box-Behnken design: an alternative for the optimization of analytical methods. *Anal Chim Acta*. 2007; 597:179–186. [PubMed: 17683728]
- Gerard, N.; Pierre, B. Process of purifying heparin, and product produced therefrom. US patent. #2989438. 1961.
- Guerrini M, Beccati D, Shriver Z, Naggi A, Viswanathan K, Bisio A, Capila I, Lansing JC, Guglieri S, Fraser B, Al-Hakim A, Gunay NS, Zhang ZQ, Robinson L, Buhse L, Nasr M, Woodcock J, Langer R, Venkataraman G, Linhardt RJ, Casu B, Torri G, Sasisekharan R. Oversulfated chondroitin sulfate is a contaminant in heparin associated with adverse clinical events. *Nature Biotechnology*. 2008; 26:669–675.
- Guerrini M, Bisio A, Torri G. Combined quantitative (1)H and (13)C nuclear magnetic resonance spectroscopy for characterization of heparin preparations. *Semin Thromb Hemost*. 2001; 27:473–482. [PubMed: 11668416]
- Jandik KA, Kruep D, Cartier M, Linhardt RJ. Accelerated stability studies of heparin. *J Pharm Sci*. 1996; 85:45–51. [PubMed: 8926582]
- Kishimoto TK, Viswanathan K, Ganguly T, Elankumaran S, Smith S, Pelzer K, Lansing JC, Sriranganathan N, Zhao GL, Galcheva-Gargova Z, Al-Hakim A, Bailey GS, Fraser B, Roy S, Rogers-Cotrone T, Buhse L, Whary M, Fox J, Nasr M, Dal Pan GJ, Shriver Z, Langer RS, Venkataraman G, Austen KF, Woodcock J, Sasisekharan R. Contaminated heparin associated with adverse clinical events and activation of the contact system. *New England Journal of Medicine*. 2008; 358:2457–2467. [PubMed: 18434646]
- Kuberan B, Beeler DL, Lech M, Wu ZL, Rosenberg RD. Chemoenzymatic synthesis of classical and non-classical anticoagulant heparan sulfate polysaccharides. *J Biol Chem*. 2003a; 278:52613–52621. [PubMed: 14519763]
- Kuberan B, Lech MZ, Beeler DL, Wu ZL, Rosenberg RD. Enzymatic synthesis of antithrombin III-binding heparan sulfate pentasaccharide. *Nat Biotechnol*. 2003b; 21:1343–1346. [PubMed: 14528313]
- Lindahl U, Li JP, Kusche-Gullberg M, Salmivirta M, Alaranta S, Veromaa T, Emeis J, Roberts I, Taylor C, Oreste P, Zoppetti G, Naggi A, Torri G, Casu B. Generation of "Neoheparin" from *E coli* K5 capsular polysaccharide. *Journal of Medicinal Chemistry*. 2005; 48:349–352. [PubMed: 15658847]
- Linhardt RJ. Heparin: an important drug enters its seventh decade. *Chemistry and Industry*. 1991; 2:45–50.
- Linhardt, Robert J.; T.T. Heparin oligosaccharides: new analogues—development and applications. In: Nieforth, ZJWaKA., editor. *Carbohydrates in Drug Design*. New York: Marcel Dekker; 1997. p. 277–341.
- Liu H, Zhang Z, Linhardt RJ. Lessons learned from the contamination of heparin. *Nat Prod Rep*. 2009; 26:313–321. [PubMed: 19240943]
- Ma C, Wang A, Qin J, Li L, Ai X, Jiang T, Tang H, Xu P. Enhanced 2,3-butanediol production by *Klebsiella pneumoniae* SDM. *Appl Microbiol Biotechnol*. 2009; 82:49–57. [PubMed: 18949476]
- Manzoni M, Bergomi S, Rollini M. Influence of the culture conditions on extracellular lyase activity related to K5 polysaccharide. *Biotechnology Letters*. 2000; 22:81–85.
- Myers, RH.; Montgomery, DC. *Response surface methodology - process and product optimization using designed experiments*. 2nd ed.. New York, NY: John Wiley & Sons; 2002.
- Okuyama, T.; Yoshida, K.; Sakurai, K.; Ogura, T.; Horie, K.; Tawada, A.; Hara, T. Method of separating and recovering mucopolysaccharides from connective tissue of animals. US patent. #3862003. 1975.
- Vidic, H-J. Process for the preparation of heparin. US patent. #4283530. 1981.
- Wang Z, Dordick JS, Linhardt RJ. *Escherichia coli* K5 heparosan fermentation and improvement by genetic engineering. *Bioengineered Bugs*. 2011a; 2:63–67. [PubMed: 21636991]
- Wang Z, Ly M, Zhang F, Zhong W, Suen A, Hickey AM, Dordick JS, Linhardt RJ. *E. coli* K5 fermentation and the preparation of heparosan, a bioengineered heparin precursor. *Biotechnol Bioeng*. 2010; 107:964–973. [PubMed: 20717972]

- Wang Z, Yang B, Zhang Z, Ly M, Takieddin M, Mousa S, Liu J, Dordick JS, Linhardt RJ. Control of the heparosan N-deacetylation leads to an improved bioengineered heparin. *Appl Microbiol Biotechnol.* 2011b; 91:91–99. [PubMed: 21484210]
- Zhang ZQ, McCallum SA, Xie J, Nieto L, Corzana F, Jimenez-Barbero J, Chen M, Liu J, Linhardt RJ. Solution structures of chemoenzymatically synthesized heparin and its precursors. *Journal of the American Chemical Society.* 2008; 130:12998–13007. [PubMed: 18767845]

Figure 1.

Conversion of heparosan to bioengineered heparin. a) Heparosan is partially *N*-deacetylated by controlled base treatment. b) Partially *N*-deacetylated heparosan is chemoselectively *N*-sulfonated with trimethylamine-sulfur trioxide complex to obtain *N*-sulfo, *N*-acetyl heparosan. c) *N*-sulfo, *N*-acetyl heparosan is enzymatically converted to bioengineered heparin through a series of enzymatic steps.

Figure 2.

A typical ^1H -NMR spectrum of the *N*-sulfo, *N*-acetyl heparosan product.

3a

3b

3c

Figure 3.

a) Response surface and corresponding contour plot of the effects of NaOH concentration and reaction time on the product *N*-acetyl content, with the reaction temperature fixed at the coded level 0 (actual level 60 °C). b) Response surface and corresponding contour plot of the effects of NaOH concentration and reaction temperature on the product *N*-acetyl content, with the reaction time fixed at the coded level 0 (actual level 3 h). c) Response surface and corresponding contour plot of the effects of reaction time and reaction temperature on the product *N*-acetyl content, with the NaOH concentration fixed at the coded level 0 (actual level 2 M).

4a

4b

4c

Figure 4.

a) Response surface and corresponding contour plot of the effects of NaOH concentration and reaction time on the product depolymerization factor, with the reaction temperature fixed at the coded level 0 (actual level 60 °C). b) Response surface and corresponding contour plot of the effects of NaOH concentration and reaction temperature on the product depolymerization factor, with the reaction time fixed at the coded level 0 (actual level 3 h). c) Response surface and corresponding contour plot of the effects of reaction time and reaction temperature on the product depolymerization factor, with the NaOH concentration fixed at the coded level 0 (actual level 2 M).

Table 1

a: Factors in actual and coded levels for the full factorial design

Factors	Symbol	Low level (−1)	High level (+1)
Heparosan concentration (g/L)	X ₁	4	10
NaOH concentration (M)	X ₂	2	10
Reaction time (h)	X ₃	3	6
Reaction temperature (°C)	X ₄	60	90

b. The full factorial design and the responses (depolymerization factor = product M_N/ starting material M_N).

Run	X ₁	X ₂	X ₃	X ₄	N-acetyl content Y ₁ (%)	Depolymerization factor Y ₂	Product recovery Y ₃ (%)
1	−1	−1	−1	−1	17.88	0.722	83.39
2	1	−1	−1	−1	17.39	0.736	92.69
3	−1	1	−1	−1	4.67	0.294	77.34
4	1	1	−1	−1	3.43	0.306	60.47
5	−1	−1	1	1	0.00	0.126	13.38
6	1	−1	−1	1	0.00	0.140	17.38
7	−1	1	−1	1	0.00	0.062	11.39
8	1	1	−1	1	0.00	0.065	4.47
9	−1	−1	1	−1	4.49	0.201	87.02
10	1	−1	1	−1	3.91	0.554	93.98
11	−1	1	1	−1	3.86	0.201	64.03
12	1	1	1	−1	3.68	0.189	48.64
13	−1	−1	1	1	0.00	0.096	7.45
14	1	−1	1	1	0.00	0.096	8.49
15	−1	1	1	1	0.00	0.061	6.75
16	1	1	1	1	0.00	0.064	4.57
17	−1	−1	−1	−1	10.82	0.557	73.47
18	1	−1	−1	−1	13.91	0.618	99.87
19	−1	1	−1	−1	3.39	0.217	65.80
20	1	1	−1	−1	3.93	0.226	57.83
21	−1	−1	−1	1	0.00	0.135	17.79

b. The full factorial design and the responses (depolymerization factor = product M_N / starting material M_N).

Run	X ₁	X ₂	X ₃	X ₄	N-acetyl content Y ₁ (%)	Depolymerization factor Y ₂	Product recovery Y ₃ (%)
22	1	-1	-1	1	0.00	0.133	17.35
23	-1	1	-1	1	0.00	0.059	4.36
24	1	1	-1	1	0.00	0.059	3.20
25	-1	-1	1	-1	3.46	0.374	64.03
26	1	-1	1	-1	3.79	0.462	84.52
27	-1	1	1	-1	6.33	0.139	48.10
28	1	1	1	-1	3.61	0.147	39.43
29	-1	-1	1	1	0.00	0.100	1.98
30	1	-1	1	1	0.00	0.100	9.17
31	-1	1	1	1	0.00	0.058	2.14
32	1	1	1	1	0.00	0.061	1.56

c. Effects of the factors from the full factorial design

N-acetyl content				Depolymerization factor				Recovery				
Variable	Coefficient	Standard error	t value	P value	Coefficient	Standard error	t value	P value	Coefficient	Standard error	t value	P value
Intercept	3.392	0.5763	5.89	0	0.23	0.01874	12.27	0	39.75	1.719	23.12	0
X ₁	−0.039	0.5763	−0.07	0.946	0.0173	0.01874	0.92	0.365	0.47	1.719	0.28	0.785
X ₂	−1.336	0.5763	−2.32	0.028	−0.0919	0.01874	−4.9	0	−8.5	1.719	−4.94	0
X ₃	−3.392	0.5763	−5.89	0	−0.0484	0.01874	−2.58	0.016	−4.01	1.719	−2.33	0.027
X ₄	−1.322	0.5763	−2.29	0.03	−0.1415	0.01874	−7.55	0	−31.54	1.719	−18.34	0

Table 2

a: Factors in actual and coded levels for the Box-Behnken design

Factors	Symbol	Coded and actual level			
		-1	0	1	
NaOH concentration (M)	X ₁	1	2	3	
Reaction time (h)	X ₂	2	3	4	
Reaction temperature (°C)	X ₃	50	60	70	

b. The Box-Behnken design and the responses

Run	X ₁	X ₂	X ₃	N-acetyl content Y ₁ (%)	Depolymerization factor Y ₂	Product recovery Y ₃ (%)
1	-1	-1	0	66.23	0.848	77.77
2	1	-1	0	8.67	0.890	76.60
3	-1	1	0	44.64	0.682	75.22
4	1	1	0	0.97	0.622	71.97
5	-1	0	-1	76.92	1.047	63.81
6	1	0	-1	18.48	0.983	73.69
7	-1	0	1	30.67	0.585	91.50
8	1	0	1	0.00	0.413	68.65
9	0	-1	-1	56.82	0.990	83.29
10	0	1	-1	32.26	0.917	83.66
11	0	-1	1	12.02	0.706	83.00
12	0	1	1	1.19	0.452	74.51
13	0	0	0	15.20	0.739	101.16
14	0	0	0	14.62	0.761	91.56
15	0	0	0	15.24	0.688	87.14

c. Regression analysis of the Box-Behnken design

N-acetyl content			Depolymerization factor						Recovery		
Variable	Coefficient	Standard error	t value	P value	Coefficient	Standard error	t value	P value	Coefficient	Standard error	P value
Intercept	15.02	1.6074	9.345	0	0.72963	0.03683	19.813	0	93.2866	4.006	0
X ₁	-23.793	0.9843	-24.172	0	-0.03175	0.02255	-1.408	0.218	-2.1743	2.453	0.416
X ₂	-8.083	0.9843	-8.212	0	-0.09502	0.02255	-4.214	0.008	-1.9135	2.453	0.471
X ₃	-17.575	0.9843	-17.855	0	-0.2227	0.02255	-9.875	0	1.6511	2.453	0.531
X ₁ ²	10.527	1.4489	7.266	0.001	0.01088	0.0332	0.328	0.756	-12.2983	3.611	0.019
X ₂ ²	4.579	1.4489	3.16	0.025	0.02028	0.0332	0.611	0.568	-5.5986	3.611	0.182
X ₃ ²	5.973	1.4489	4.122	0.009	0.01632	0.0332	0.491	0.644	-6.573	3.611	0.128
X ₁ X ₂	3.472	1.3921	2.494	0.055	-0.0257	0.03189	-0.806	0.457	-0.5214	3.469	0.886
X ₁ X ₃	6.941	1.3921	4.986	0.004	-0.02684	0.03189	-0.842	0.438	-8.1826	3.469	0.065
X ₂ X ₃	3.433	1.3921	2.466	0.057	-0.04509	0.03189	-1.414	0.217	-2.2155	3.469	0.551
R ² = 0.995			R ² = 0.960						R ² = 0.823		